UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A Rational Trade-Off Between the Costs and Benefits of Automatic and Controlled Processing

Permalink

https://escholarship.org/uc/item/5qs4b625

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Mittenbühler, Maximilian
Schwöbel, Sarah
Dignath, David
et al.

Publication Date

2024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A Rational Trade-Off Between the Costs and Benefits of Automatic and Controlled Processing

Maximilian Mittenbühler[1] **(maximilian.mittenbuehler@uni-tuebingen.de),**
Sarah Schwöbel[2,3] **(sarah.schwoebel@tu-dresden.de), David Dignath**[4] **(david.dignath@uni-tuebingen.de),**
Stefan J. Kiebel[2,3] **(stefan.kiebel@tu-dresden.de), & Martin V. Butz**[1,4] **(martin.butz@uni-tuebingen.de)**
[1]Department of Computer Science, University of Tübingen, Tübingen, Germany
[2]Department of Psychology, TUD Dresden University of Technology, Dresden, Germany
[3]Centre for Tactile Internet with Human-in-the-Loop (CeTI), TUD Dresden University of Technology, Dresden, Germany
[4]Department of Psychology, University of Tübingen, Tübingen, Germany

Abstract

Humans seem to arbitrate between automatic and controlled processing by optimizing a trade-off between cognitive effort and performance. Previous research has described ways of how these costs and benefits can be quantified and how the trade-off between them can be performed. However, it remains unclear how the costs should be weighed relative to the benefits and how the cost of the arbitration mechanism itself factors in. Here, we derive measures for these separate factors from a single objective: the variational free energy. We demonstrate that by minimizing this objective, the trade-off between automatic and controlled processing as well as meta-control is optimized implicitly. As a proof of concept, we show that the congruency and proportion congruency effects in the Stroop task directly result from this optimization, given an environment with specific statistical regularities.

Keywords: conflict tasks, meta-control, free energy principle, variational inference, resource rationality

Introduction

Rational behavior describes decision making based on predicted outcomes and their utility (Von Neumann & Morgenstern, 1944), a type of decision making that is also a defining feature of controlled processing (Evans & Stanovich, 2013). In contrast, theories of resource rationality argue that it is often not rational to fully rely on controlled processing because of the cognitive effort it involves (Gershman, Horvitz, & Tenenbaum, 2015; Griffiths, Lieder, & Goodman, 2015; Lewis, Howes, & Singh, 2014; Ortega & Braun, 2013; Shenhav et al., 2017). Instead, it may be rational in some situations to rely on automatic processing, i.e., to make decisions based on previously learned stimulus-response (S-R) associations (Evans & Stanovich, 2013). This necessitates a mechanism that arbitrates between these two modes of processing, referred to as meta-control (Eppinger, Goschke, & Musslick, 2021). It has been proposed that this mechanism reflects a cost-benefit analysis, weighing the increase in flexibility and accuracy of controlled processing against the reduction in cognitive effort by using automatic processing (Kool, Gershman, & Cushman, 2017; Shenhav, Botvinick, & Cohen, 2013). The benefits of control have been quantified by the expected increase in reward, whereas its costs have been expressed as either a simple function of the control signal strength (Bustamante, Lieder, Musslick, Shenhav, & Cohen, 2021; Shenhav et al., 2013; Verguts, Vassena, & Silvetti, 2015) or the informational cost accrued by integrating the task goal in the belief over actions (Butz, 2022; Ortega & Braun, 2013; Parr, Holmes, Friston, & Pezzulo, 2023; Zénon, Solopchuk, & Pezzulo, 2019). These formalizations, however, leave two challenges unresolved: Firstly, they require a parameter to scale the trade-off between the costs and benefits, hampering precise resource rational predictions. Secondly, they imply that the cognitive system can only optimize this trade-off (perform meta-control) by explicitly estimating the costs and benefits of control and weighing them against each other. How these operations are performed and what additional costs they incur remains unclear.

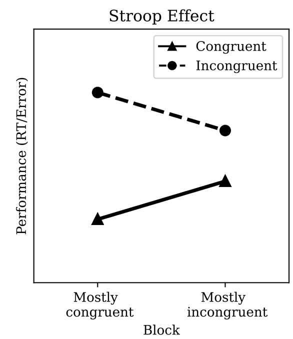

Figure 1: Schematic illustration of congruency effect (main effect of congruency) and PC effect (interaction effect) in the Stroop task for both reaction time (RT) and error rate. See, e.g., Spinelli and Lupker (2023).

Here, we address these issues by deriving the costs and benefits of automatic and controlled processing as well as meta-control from a single objective function, the variational free energy (Parr, Pezzulo, & Friston, 2022; Schwöbel, Kiebel, & Marković, 2018). We show that by minimizing the free energy of a generative model, a cognitive system implicitly optimizes these costs and benefits. Our formalization provides a straightforward rationale for the existence this arbitration mechanism in humans and allows to make specific predictions about when and how to utilise it rationally.

As a proof of concept, we draw on two behavioral effects from the classical Stroop task (Stroop, 1935), which have been extensively studied to measure automatic and controlled processing in humans (see Chuderski & Smolen, 2016). In the Stroop task, participants must name the print color of a carrier word while ignoring the color meaning of the car-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

rier word. As illustrated in Figure 1, participants respond more slowly and less accurately in incongruent trials (i.e., trials with a mismatch between print color and word meaning) compared to congruent trials. This *congruency effect* demonstrates automatic processing in human decision making. Furthermore, the congruency effect is larger in blocks that contain mostly congruent trials and vice versa, which is known as the *proportion congruency (PC) effect* (Spinelli & Lupker, 2023). This effect arguably shows to what extent humans modulate the balance between automatic and controlled processing. We contribute to the many existing computational models predicting these effects (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Chuderski & Smolen, 2016; Grahek, Musslick, & Shenhav, 2020; Jiang, Heller, & Egner, 2014) by demonstrating that the effects are a rational consequence of minimizing free energy in environments with particular statistical regularities.

Methods

Model

We assume that a human's generative model of the Stroop task can be described as a Bayesian network (Figure 2), consisting of six categorical variables: the stimulus $s_1 \in \{\text{RED}, \text{RED}, \text{GREEN}, \text{GREEN}\}$, the action $a \in \{\text{red, green}\}$, the outcome state $s_2 \in \{\text{correct, incorrect}\}$, the associated observation $o \in \{\text{good, bad}\}$ as well as the current context $c_t \in \{1, 2\}$. Here, the first context, which we refer to as the *automatic context*, defines an informative S-R mapping given by the parameter ϕ, and the second context, which we refer to as the *non-automatic context*, defines a uniform S-R mapping.

Corresponding to Figure 2, the generative model can be expressed mathematically as

$$p(c_t, a, s_2, o | s_1, \phi) = p(c_t) p(a | c_t, s_1, \phi) \times p(s_2 | s_1, a) p(o | s_2) \quad (1)$$

Crucially, we extend this generative model by multiplying it with a prior $p(o)$ that scores the utility of future observations. This method is frequently used in other Bayesian decision making or active inference works to define the agent's preferences or goal, enabling value-based decision making and therefore controlled processing (Friston et al., 2015; Parr et al., 2023; Schwöbel, Marković, Smolka, & Kiebel, 2021; Solway & Botvinick, 2012).

We start by giving a conceptual overview of the model before detailing the inference process more formally in the following section. To behave successfully, the agent has to change its beliefs over the hidden variables in the model in light of new evidence $p(s_1)$ and its preferences $p(o)$. Therefore, there are two sources of information influencing the belief over actions. To incorporate $p(o)$ (red edges in Figure 2), the agent has to invert its generative model in order to infer which state most likely causes high utility observations ($p(o|s_2)$) and, consecutively, which action it must perform to reach that state ($p(s_2|s_1, a)$). This is an instantiation of controlled processing. We fix $p(o) = (.95, .05)$, meaning that the agent highly prefers "good" over "bad" observations. To incorporate $p(s_1)$ (blue edge in Figure 2), the agent can use the direct mapping from stimulus to action (S-R mapping, $p(a|c_t, s_1, \phi)$), yielding a prior over actions. We define this as an automatic process because it is independent of the agent's goal and it does not require the generative model to be inverted (see Discussion). The parameter ϕ determines how strong this mapping is by specifying the probability of an action given a particular stimulus (for the automatic context). According to Bayes' rule, the automatic and controlled processes are combined depending on their relative uncertainty (see also Schwöbel et al., 2021). This means that the more information the automatic process contains (i.e., the more the prior diverges from a uniform distribution), the stronger its influence on the inference process.

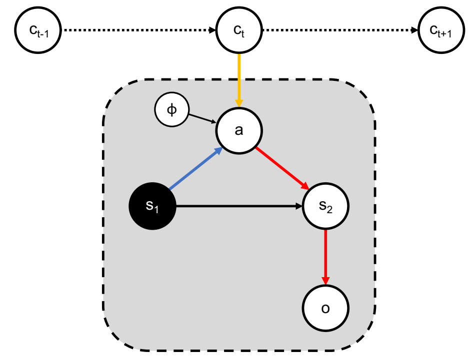

Figure 2: Probabilistic generative model of the Stroop effect depicted as a Bayesian graph, which maps to the generative model in Eq. 1. The information processes that are unfolding within a trial are shown in the gray area. Observed nodes have black background. Red edges: controlled processing; blue edge: automatic processing; yellow edge: meta-control; dotted edges: context transitions across trials t.

Crucially, the strength of the S-R mapping (and therefore the balance between automatic and controlled processing) is also context-dependent (yellow edge in Figure 2). This is because the mapping is informative in the automatic context but uniform in the non-automatic context, and thus changes depending on the belief over contexts $p(c_t)$. As depicted by the dotted edges in Figure 2, information about which context the agent expects is carried over from the previous to the current trial t. After having inferred a posterior over actions, the agent can invert the dependency between c_t and a to infer a posterior belief over the current context. Specifically, if the posterior over actions corresponds to the prior over actions (i.e., the automatic process was "helpful"), the probability of the automatic context is increased, and decreased otherwise. This information is, again, passed on to the next trial, resulting in a continuous meta-control process. We use a parameter, γ (see Equation 8), that governs how much the belief over contexts is changed, and therefore effectively determines the flexibility of meta-control.

Variational inference

Based on the free energy principle (Parr et al., 2022), we assume that the agent performs variational inference to obtain a joint posterior q by minimizing the variational free energy $F[q]$.

$$F[q] = D_{KL}[q(c_t, a, s_2, o|\phi)||p(c_t, a, s_2, o|s_1, \phi)p(o)] \quad (2)$$

The free energy essentially measures the KL-divergence between the model (right part of the KL-divergence), which includes the agent's preferences, and the variational posterior (left part of the KL-divergence). It is an information-theoretic quantity measured in *natural units of information* (nats). Note that our formalization differs from the *expected free energy* (e.g., Parr et al., 2023) as it contains no additional term for expected information gain, which plays no role in our scenario (for a comparison see Millidge, Tschantz, & Buckley, 2021). Importantly however, our formalization follows directly from variational inference (Schwöbel et al., 2018). Furthermore, be aware that the posterior q does not explicitly depend on s_1 because it is the result of an optimization process, but it does incorporate this observation.

To be able to interpret this quantity and relate it to the costs and benefits of automatic and controlled processing, we decompose it in the following. We posit that the agent minimizes free energy on two different levels—or time-scales—one within each trial and one across trials (cf., Figure 2). Between these levels, we assume the mean field approximation (Friston et al., 2015), which yields

$$F[q] = \underbrace{D_{KL}[q(c_t)||p(c_t)]}_{\text{cognitive effort (between)}} + \underbrace{\sum_{c_t} q(c_t) F[q|c_t]}_{\text{free energy (within)}}, \quad (3)$$

where

$$F[q|c_t] = D_{KL}[q(a, s_2, o|c_t, \phi)||p(a, s_2, o|c_t, s_1, \phi)p(o)] \quad (4)$$

Let us first consider how free energy is minimized within a trial t. On this faster time-scale, we assume that the agent cannot adjust its belief over contexts, i.e., $q(c_t) = p(c_t)$. Observing a stimulus s_1 then automatically produces a prior over actions $p(a|c_t, s_1, \phi)$ (automatic processing). For controlled processing, the agent has to approximate a posterior over actions by minimizing $F[q|c_t]$. Given the dependency structure defined by the generative model (Figure 2 and Equation 1) and assuming the Bethe approximation within a trial (Schwöbel et al., 2018; Yedidia, Freeman, & Weiss, 2003), we can decompose $F[q|c_t]$ as follows:

$$F[q|c_t] = \underbrace{D_{KL}[q(a, s_2|c_t, \phi)||p(a, s_2|c_t, s_1, \phi)]}_{\text{cognitive effort (within)}} \quad (5)$$

$$+ \underbrace{D_{KL}[q(s_2, o|c_t)||p(o|s_2)p(o)]}_{\text{divergence from goal}} \quad (6)$$

$$+ \underbrace{H(q(s_2|c_t))}_{\text{state uncertainty}} \quad (7)$$

It follows that minimizing free energy within a trial can be understood as a trade-off between three quantities. The first is a KL-divergence between the prior and the posterior over the latent variables, measuring how many nats have to be spent to change the agent's belief. We refer to this as the cognitive effort involved in making a decision within a trial. Since we assume that the prior results from automatic processing, it is clear that automatic processing can reduce cognitive effort (if the prior is similar to the posterior). The second term is a KL-divergence between the prior over outcomes (the agent's goal) with the corresponding state and the posterior over outcomes and states. This quantifies by how many nats the expected states and outcomes of the agent's action diverge from its desired states and outcomes, or in other words, how well the agent expects to perform. The third term is the entropy of the agent's belief over future states, measuring in nats how uncertain the agent is to reach these states. To perform variational inference therefore means to find a posterior belief that optimizes this trade-off. In our implementation, we obtain this posterior by performing belief propagation, which has been shown to be equivalent to minimizing free energy under the Bethe approximation (Yedidia et al., 2003). After having obtained the posterior, the agent samples an action which it then executes. Therefore, the agent's expected goal divergence directly translates to the agent's real accuracy on average.

Between two trials, we assume that the agent has already minimized $F[q|c_t]$ (and executed an action), but can adjust its belief over contexts (meta-control). Turning back to Equation 3, we see that, at this point, the agent performs another cost-benefit analysis, trading off the cost of changing its belief over contexts against the resulting reduction in within-trial free energy. In other words, the agent has to weigh the cost of meta-control (changing an additional belief) against the potential improvements for action selection (within-trial free energy).

Assuming $F[q|c_t]$ to be fixed, Equation 3 can be transformed into an update equation for the belief over contexts (see Schwöbel et al., 2018):

$$q(c_t) = \frac{p(c_t)e^{-\gamma F[q|c_t]}}{\sum_{c'_t} p(c'_t)e^{-\gamma F[q|c'_t]}} \quad (8)$$

This shows that the belief is changed in favor of contexts that are associated with a smaller within-trial free energy. For the present model, this means that if the prior over actions specified by a context was helpful (i.e., similar to the posterior), this context subsequently receives a higher probability. Here, we follow previous work in the active inference and reinforcement learning literature (Friston et al., 2015; Ortega & Braun, 2013; Schwöbel et al., 2021; Verguts et al., 2015) by utilizing a parameter γ to determine the extent to which the context is updated (i.e., the flexibility of meta-control).

Information about which context should be expected is retained over time through

$$p(c_t) = \sum_{c_{t-1}} p(c_t|c_{t-1}) q(c_{t-1}) \quad (9)$$

We fix the off-diagonal elements in $p(c_t|c_{t-1})$ to .05, resulting in a slight blurring of the context belief over time, which we found useful to keep the model flexible enough to incorporate environmental changes. Note that the off-diagonal elements in $p(s_2|s_1,a)$ and $p(o|s_2)$ are set to .01, making them nearly deterministic.

Exploiting statistical structure in the environment

We assume that humans further minimize free energy on an even longer time scale by learning the strength of S-R mappings and of their adjustment (i.e., of meta-control). In our formalization, this can be modeled by optimizing the parameters ϕ and γ, respectively. As a proxy for the learning process we perform a grid search over the parameter space, minimizing the average free energy over an agent's life time.

$$\phi^*, \gamma^* = \underset{\phi,\gamma}{\arg\min}\, F[q] \tag{10}$$

The optimal strength of S-R mappings and meta-control depends on the statistical structure of an environment. To illustrate this, we assume that an agent interacts over its life time with an environment that consists of two contexts (not to be confused with the contexts that are part of the agent's generative model). In the first context, the word the agent reads is associated with the action it executes in 90% of the cases, introducing a strong contingency between word meaning and action. In the second context this contingency is instead invalid in 90% of the cases. We further assume that an agent switches back and forth between these contexts but spends more time in the first (75% of interactions). In summary, this means that there is relationship between word meaning and action of 70% across contexts, but that this relationship is even stronger in the first context (90%) and much weaker in the second context (10%).

To analyze the effects of the agents' learning abilities, we compare three setups: a first "controlled processing agent" does not learn structure at all ($\phi = .5$ and $\gamma = 0$), a second "mixed processing agent" learns a general S-R association (ϕ is optimized, $\gamma = 0$), and a third "meta-control agent" additionally learns to change its belief over contexts (both ϕ and γ are optimized).

Simulating the Stroop task

Once learning in the environment described above is complete, we freeze the agents' parameter values and subject them to the Stroop task. The task consists of two blocks with 400 trials each, where one block contains 25% congruent trials (mostly incongruent) and one contains 75% congruent trials (mostly congruent). To obtain RT predictions, we assume that changing a belief happens at a constant rate and therefore

$$RT = NDT + \alpha \cdot D_{KL}[q(a,s_2|c_t,\phi)||p(a,s_2|c_t,s_1,\phi)] \tag{11}$$

For illustration purposes we choose a change rate (α) of 0.005 nats per millisecond with a non-decision time (NDT) of 650 ms.

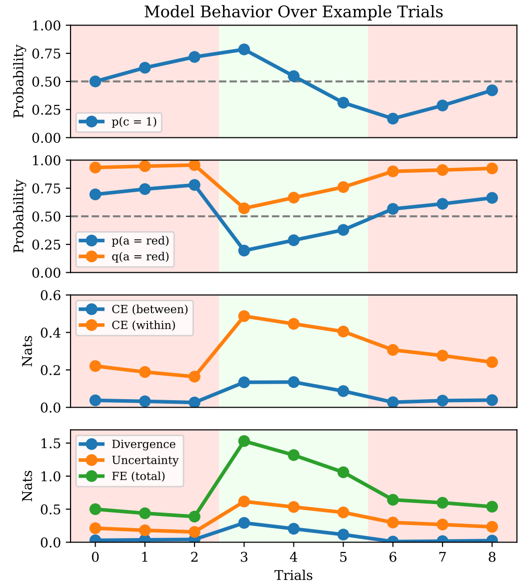

Figure 3: Congruent trials with s_1 = RED are shaded red, incongruent trials with s_1 = GREEN are shaded green. 1: Agent's belief over contexts at the beginning of a trial (dashed line marks uniform probability). 2: Agent's prior (blue line) and posterior belief (orange line) over actions. 3: Cognitive effort (CE) dispensed for controlled processing (CE within) and context adaptation (CE between). 4: Information-theoretic measures for divergence from goal, state uncertainty, and total free energy. Parameters are set to optimal values $\gamma^* = 1.21$, $\phi^* = 0.89$.

Results

Conceptual illustration

As an example, we show in Figure 3 the model's behavior and the associated information-theoretic quantities for a sequence of nine trials. In this specific example, the first action (a = red) is correct in all nine trials as the word color is always red. First, the agent is presented with three congruent trials (s_1 = RED), i.e., the automatic response tendency is helpful. The agent starts with a uniform belief over contexts, leading to a moderate automatic response tendency ($p(a = \text{red}|c_t, s_1 = \text{RED}, \phi) > .5$). Because the response tendency is helpful three times in a row, the agent increasingly beliefs to be in the automatic context ($p(c = 1)$) and therefore relies more strongly on this response tendency in the next trial. This leads to a decrease in cognitive effort for action selection (within) and a decrease in the overall free energy. However, when the agent is presented with a series of three incongruent trials (s_1 = GREEN), the response tendency is misleading, which is accompanied with a substantial increase in cognitive effort for action selection (within) and overall free energy. To counteract this, more cognitive effort (between) is spent to adjust the current belief over contexts in favor of the non-automatic context. The reverse happens

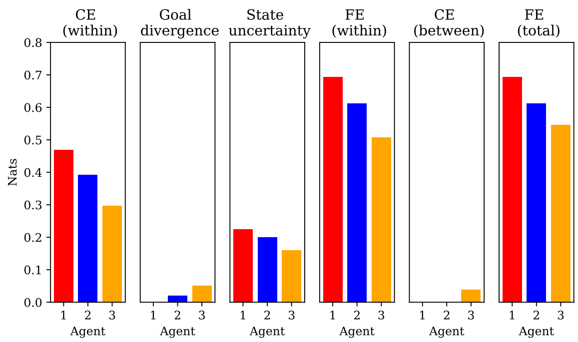

Figure 4: Life-time average of information-theoretic quantities for the three agents (1: controlled processing; 2: mixed processing; 3: meta-control). CE: cognitive effort. FE: free energy.

when the agent is again presented with a series of congruent trials.

Rational automatic processing and meta-control

Figure 4 shows the life-time average of the information-theoretic quantities defined above for the three agents. For the "mixed control agent", $\phi^* = 0.73$ was optimal and the figure demonstrates that learning this S-R association reduces the overall free energy by about 12% in the example environment. This reduction is the result of trading off a reduced cognitive effort and state uncertainty against an increased divergence from the goal (which directly translates to the agent's accuracy, see Methods section). Therefore, relying on automatic processing is rational in this environment.

For the "meta-control agent", $\gamma^* = 1.21$ and $\phi^* = 0.89$ was found to be optimal, suggesting that through the meta-control mechanism, the agent was able to utilize a stronger S-R association. Figure 4 shows that this reduces the overall free energy by 19% compared to learning no S-R mapping at all and by 9% compared to learning a general S-R mapping. This is because the cognitive effort involved in controlled processing is reduced substantially, at a relatively small cost of an increased error rate and cognitive effort (between) for updating the context belief.

Figure 5 shows the same information-theoretic quantities depending on the values of γ and ϕ for the "meta-control agent". In the top row it can be seen that, as the S-R mapping becomes stronger, the cognitive effort needed for action selection and the uncertainty of future state predictions is reduced. However, the divergence from the agent's goal is increased. The lower left graph shows that when doing the cost-benefit-analysis, the agent does best if it accepts to stray slightly from its goal to save cognitive effort and make more precise predictions. This means that, in this particular environment, the agent benefits from relying on automatic processing to a larger extent.

The same graph also suggest that the stronger the meta-control, the better—arguably because this leads to a more fine-grained balance between automatic and controlled processing. However, when taking the cost of meta-control (lower middle graph) into account, the lower right graph shows that the agent does best if it has a strong automatic response tendency in one context ($\phi = .89$), and if it weighs the costs of meta-control roughly equally against the benefits ($\gamma = 1.21$). It can also be seen that a strong S-R association is only beneficial for the agent if it has the right amount of meta-control—both too weak and too strong settings of γ lead to considerable increases in free energy. Thus, the graph defines precisely how to best arbitrate between automatic and controlled processing, given this environment.

Resulting effects in the Stroop task

Figure 6 shows that the "controlled processing agent" does not exhibit a congruency or proportion congruency (PC) effect. The "mixed processing agent" shows the congruency but not the PC effect. Only the "meta-control agent", which fully minimizes free energy by exploiting structure in the environment, shows both a congruency and a PC effect and is therefore qualitatively most similar to human behavior (compare with Figure 1).

Discussion

We found that (1) measures of the costs and benefits of automatic and controlled processing as well as meta-control can be derived from the free energy principle, (2) relying on automatic processing can be rational depending on the environment, (3) using a specific amount of meta-control is rational depending on the environment, and (4) the congruency and PC effect in the Stroop task are a direct consequence of minimizing free energy in the given environment.

Our derived measures align well with previous formalizations, such as the expected value of control (Griffiths et al., 2019; Shenhav et al., 2013; Bustamante et al., 2021; Lieder, Shenhav, Musslick, & Griffiths, 2018) or the utility of effort (Verguts et al., 2015). Moreover, they are partially equivalent to other information-theoretic approaches (Butz, 2022; Parr et al., 2023; Zénon et al., 2019). We extend these previous measures by including meta-control (see also Bustamante et al., 2021) and demonstrate that all measures are implicitly optimized by minimizing free energy. This is particularly relevant in the light of previous research that has investigated how these measures can be estimated efficiently, e.g., through learning associations between stimuli and the value of control (Lieder et al., 2018; Bustamante et al., 2021). It also sheds further light on the "homunculus problem" of cognitive control (i.e., "who controls the controller?"; Botvinick & Cohen, 2014) as it defines an optimum for both the use of controlled processing and meta-control itself (Boureau, Sokol-Hessner, & Daw, 2015).

The present model differs from previous influential works which have assumed that cognitive control adjusts the influence of competing automatic processes (S-R mappings) on the response (e.g., Botvinick et al., 2001). We instead assume that automatic processes provide a prior belief over actions

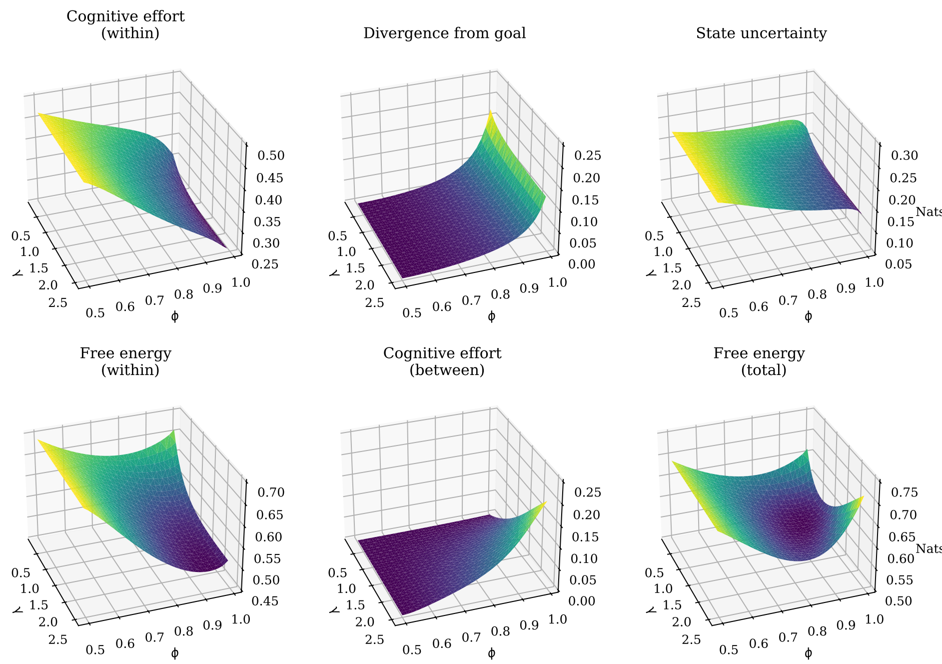

Figure 5: Life-time average of information-theoretic quantities for the "meta-control agent". The two horizontal axes show parameters controlling the strength of the S-R mapping (ϕ) and the context update (γ). The vertical axis and color correspond to respective values in nats (blue: lower; yellow: higher).

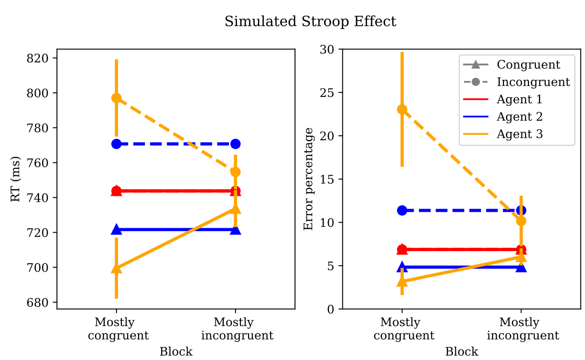

Figure 6: Predictions of RT and accuracy effects in the Stroop task for the three agents (1: controlled processing; 2: mixed processing; 3: meta-control). Compare with Figure 1. The same colors as in Figure 4 are used to differentiate the three agents. Error bars depict one SD.

and that controlled processing updates this belief by conditioning it on desired future observations (see also Schwöbel et al., 2021). While this makes the arbitration between them straightforward, it also describes them as two fundamentally different operations: The former comprises a simple mapping from stimulus to response, whereas the latter requires an inversion of the generative model. S-R mappings can be learned and therefore be executed with little effort, for instance, as a forward pass through a neural network. Arguably, however, learning a simple mapping for the inversion process (see, e.g., Kingma & Welling, 2022) seems impractical both because the learning problem is much harder and because it would nullify the capacity to react flexibly to changes in the environment (e.g., contingency degradation). This provides a mechanistic argument as to why the use of cognitive control is experienced as effortful (Kool & Botvinick, 2018) and is therefore sometimes substituted with simpler automatic processes.

Furthermore, the free energy objective urges the agent to minimize cognitive effort even though we do not define a limit on the amount of effort that can be exerted or the time available to do so. Arguments about metabolic resource limitations or opportunity costs (see, e.g., Kool & Botvinick, 2018), therefore, have no bearing on it. It instead suggests that constraining cognitive effort in itself might be adaptive, for instance, because it reduces the risk of overfitting to data—a common notion in statistical modeling. In our case, this means that the agent is biased by its previous experience (i.e., automatic processing) instead of only relying on the immediately available sensory observations (i.e., controlled processing)—an arguably adaptive approach. Similar to this, previous work has argued that cognitive control might be limited not because of an inherent scarcity of cognitive resources but because the constraint optimizes a trade-off between cognitive stability and flexibility (Kool & Botvinick, 2018; Musslick & Cohen, 2021). Our results corroborate this account, considering that the limitation is effectively implemented by the cognitive effort discounting the controlled process.

Acknowledgements

The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Maximilian Mittenbühler.

This work was partially supported by the German Research Foundation (DFG) within research unit FOR 2718 - project "Interplay of Amodal and Modal Encodings Underlying Directional Space-Metric Associations" (381713393) as well as by the Machine Learning Cluster of Excellence, under Germany's Excellence Strategy – EXC number 2064/1 – Project number 390727645.

This work was also partially supported by the German Research Foundation (DFG) via SFB 940 — Project number 178833530 and as part of Germany's Excellence Strategy—EXC 2050/1 — Project number 390696704 — Cluster of Excellence, Centre for Tactile Internet with Human-in-the-Loop (CeTI) of TUD Dresden University of Technology, and by the Saxon State Ministry of Science, Culture and Tourism (SMWK) as part of support for profile-defining structural units of the TUD in 2023/24.

References

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. doi: 10.1037/0033-295X.108.3.624

Botvinick, M. M., & Cohen, J. D. (2014, August). The Computational and Neural Basis of Cognitive Control: Charted Territory and New Frontiers. *Cognitive Science*, *38*(6), 1249–1285. doi: 10.1111/cogs.12126

Boureau, Y.-L., Sokol-Hessner, P., & Daw, N. D. (2015, November). Deciding How To Decide: Self-Control and Meta-Decision Making. *Trends in Cognitive Sciences*, *19*(11), 700–710. doi: 10.1016/j.tics.2015.08.013

Bustamante, L., Lieder, F., Musslick, S., Shenhav, A., & Cohen, J. (2021, June). Learning to Overexert Cognitive Control in a Stroop Task. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(3), 453–471. doi: 10.3758/s13415-020-00845-x

Butz, M. V. (2022, June). Resourceful Event-Predictive Inference: The Nature of Cognitive Effort. *Frontiers in Psychology*, *13*, 867328. doi: 10.3389/fpsyg.2022.867328

Chuderski, A., & Smolen, T. (2016). An integrated utility-based model of conflict evaluation and resolution in the Stroop task. *Psychological Review*, *123*(3), 255–290. doi: 10.1037/a0039979

Eppinger, B., Goschke, T., & Musslick, S. (2021, June). Meta-control: From psychology to computational neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(3), 447–452. doi: 10.3758/s13415-021-00919-4

Evans, J. S. B. T., & Stanovich, K. E. (2013, May). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. doi: 10.1177/1745691612460685

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015, October). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214. doi: 10.1080/17588928.2015.1020053

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015, July). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. doi: 10.1126/science.aac6076

Grahek, I., Musslick, S., & Shenhav, A. (2020, May). A computational perspective on the roles of affect in cognitive control. *International Journal of Psychophysiology*, *151*, 25–34. doi: 10.1016/j.ijpsycho.2020.02.001

Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019, October). Doing more with less: meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, *29*, 24–30. doi: 10.1016/j.cobeha.2019.01.005

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015, April). Rational Use of Cognitive Resources: Levels of Analysis Between the Computational and the Algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. doi: 10.1111/tops.12142

Jiang, J., Heller, K., & Egner, T. (2014, October). Bayesian modeling of flexible cognitive control. *Neuroscience & Biobehavioral Reviews*, *46*, 30–43. doi: 10.1016/j.neubiorev.2014.06.001

Kingma, D. P., & Welling, M. (2022, December). *Auto-Encoding Variational Bayes.* arXiv. (arXiv:1312.6114 [cs, stat])

Kool, W., & Botvinick, M. (2018, December). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. doi: 10.1038/s41562-018-0401-9

Kool, W., Gershman, S. J., & Cushman, F. A. (2017, September). Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychological Science*, *28*(9), 1321–1333. doi: 10.1177/0956797617708288

Lewis, R. L., Howes, A., & Singh, S. (2014, April). Computational Rationality: Linking Mechanism and Behavior Through Bounded Utility Maximization. *Topics in Cognitive Science*, *6*(2), 279–311. doi: 10.1111/tops.12086

Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018, April). Rational metareasoning and the plasticity of cognitive control. *PLOS Computational Biology*, *14*(4), e1006043. doi: 10.1371/journal.pcbi.1006043

Millidge, B., Tschantz, A., & Buckley, C. L. (2021, Febru-

ary). Whence the Expected Free Energy? *Neural Computation*, *33*(2), 447–482.
doi: 10.1162/neco$_{a0}$1354

Musslick, S., & Cohen, J. D. (2021, September). Rationalizing constraints on the capacity for cognitive control. *Trends in Cognitive Sciences*, *25*(9), 757–775.
doi: 10.1016/j.tics.2021.06.001

Ortega, P. A., & Braun, D. A. (2013, May). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2153), 20120683.
doi: 10.1098/rspa.2012.0683

Parr, T., Holmes, E., Friston, K. J., & Pezzulo, G. (2023, June). Cognitive effort and active inference. *Neuropsychologia*, *184*, 108562.
doi: 10.1016/j.neuropsychologia.2023.108562

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: the free energy principle in mind, brain, and behavior*. Cambridge, Massachusetts: The MIT Press.

Schwöbel, S., Kiebel, S. J., & Marković, D. (2018). Active Inference, Belief Propagation, and the Bethe Approximation. *Neural Computation*, *30*, 2530–2567. doi: doi:10.1162/neco$_{a0}$1108

Schwöbel, S., Marković, D., Smolka, M. N., & Kiebel, S. J. (2021, February). Balancing control: A Bayesian interpretation of habitual and goal-directed behavior. *Journal of Mathematical Psychology*, *100*, 102472.
doi: 10.1016/j.jmp.2020.102472

Shenhav, A., Botvinick, M., & Cohen, J. (2013, July). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron*, *79*(2), 217–240.
doi: 10.1016/j.neuron.2013.07.007

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017, July). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*(1), 99–124.
doi: 10.1146/annurev-neuro-072116-031526

Solway, A., & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. *Psychological Review*, *119*(1), 120–154.
doi: 10.1037/a0026435

Spinelli, G., & Lupker, S. J. (2023, May). Robust evidence for proactive conflict adaptation in the proportion-congruent paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(5), 675–700.
doi: 10.1037/xlm0001144

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643–662.

Verguts, T., Vassena, E., & Silvetti, M. (2015, March). Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. *Frontiers in Behavioral Neuroscience*, *9*.
doi: 10.3389/fnbeh.2015.00057

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Yedidia, J. S., Freeman, W. T., & Weiss, Y. (2003). Understanding Belief Propagation and its Generalizations. *Exploring artificial intelligence in the new millennium*, *8*, 239–269.

Zénon, A., Solopchuk, O., & Pezzulo, G. (2019, February). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, *123*, 5–18.
doi: 10.1016/j.neuropsychologia.2018.09.013